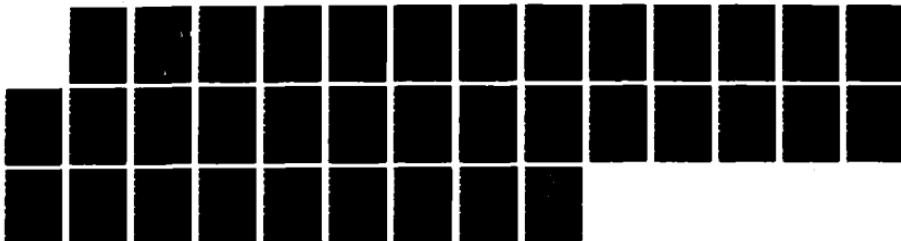
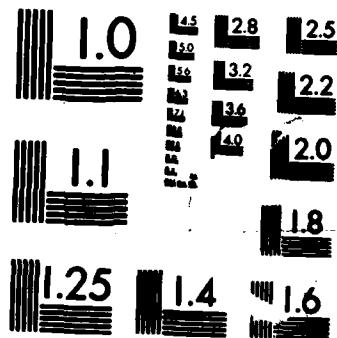


ND-H187 568 THE CONVERGENCE RATE FOR THE STRONG LAW OF LARGE  
171  
NUMBERS: GENERAL LATTICE DISTRIBUTIONS(U) STANFORD UNIV  
CA DEPT OF STATISTICS J A FILL ET AL OCT 87 TR-3  
UNCLASSIFIED N08814-87-K-8878 F/G 12/3 NL





MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

OTC FILE COPY

1

AD-A187 560

THE CONVERGENCE RATE FOR THE STRONG LAW OF  
LARGE NUMBERS: GENERAL LATTICE DISTRIBUTIONS

by

James Allen Fill and Michael J. Wichura  
Stanford University and The University of Chicago

TECHNICAL REPORT NO. 3  
OCTOBER 1987

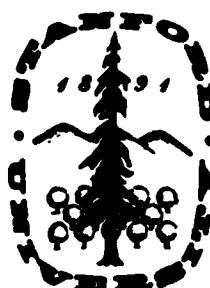
DTIC  
ELECTED  
NOV 10 1987  
S CSD

PREPARED UNDER CONTRACT  
N00014-87-K-0078 (NR-042-373)  
FOR THE OFFICE OF NAVAL RESEARCH

Reproduction in Whole or in Part is Permitted  
for any Purpose of the United States Government

Approved for public release; distribution unlimited

DEPARTMENT OF STATISTICS  
STANFORD UNIVERSITY  
STANFORD, CALIFORNIA



87 10 23 097

**THE CONVERGENCE RATE FOR THE STRONG LAW OF  
LARGE NUMBERS: GENERAL LATTICE DISTRIBUTIONS**

by

**James Allen Fill and Michael J. Wichura**  
**Stanford University and The University of Chicago**

**TECHNICAL REPORT NO. 3**

**OCTOBER 1987**

**PREPARED UNDER CONTRACT  
N00014-87-K-0078 (NR-042-373)  
FOR THE OFFICE OF NAVAL RESEARCH**

**Reproduction in Whole or in Part is Permitted  
for any Purpose of the United States Government**

**Approved for public release; distribution unlimited**



**DEPARTMENT OF STATISTICS  
STANFORD UNIVERSITY  
STANFORD, CALIFORNIA**

Account for	
NTIS CRAGI	<input checked="" type="checkbox"/>
DTC TAB	<input type="checkbox"/>
Unpublished	<input type="checkbox"/>
Intellectual	<input type="checkbox"/>
By	
Dist. Factor	
A. Mobility, Function	
Ref. No.	
Oral	<input type="checkbox"/>
7. Oral and/or Special	
A-1	

## 1. Introduction and formulation of results.

Let  $X_1, X_2, \dots$  be a sequence of independent random variables with common distribution function  $F$ . Denote the random walk of partial sums by  $S_n = \sum_{k=1}^n X_k$ , with  $S_0 \equiv 0$ . To facilitate notation we introduce another random variable  $X$  distributed according to  $F$ . Throughout this paper we adopt

$$\text{Assumption 0: } EX = 0, \quad \text{Var } X = 1$$

and write

$$K(\xi) = \log E(e^{\xi X})$$

for the cumulant generating function (cgf) of  $X$ ,  $I$  for the set of  $\xi$  for which  $K(\xi) < \infty$ , and  $I^0$  for its interior. Let  $\epsilon (> 0$  throughout) and  $\alpha$  be given real numbers, and denote by  $g$  the straight line

$$g(t) = \alpha + \epsilon t.$$

The notation  $E(Y; A)$  used herein is shorthand for  $E(YI_A)$ .

The problem addressed in this paper is to determine the asymptotic rate at which the boundary crossing probabilities

$$(1.1) \quad P\{S_m > g(m)\}$$

and

$$(1.2) \quad p_m = P\{S_n > g(n) \text{ for some } n \geq m\}$$

converge to zero as  $m \rightarrow \infty$ . These probabilities are of interest largely because when  $\alpha = 0$  the weak and strong laws of large numbers (WLLN and SLLN) amount, respectively, to the assertions that (1.1) and (1.2) tend to zero. Fill (1980, 1983) has also used approximating linear boundaries with nonzero intercept to obtain an asymptotic expression for (1.2) when  $g$  is a curved boundary.

Fill (1980, theorem 3.4.1), slightly generalizing computations of Bahadur and Ranga Rao (1960) and Siegmund (1975), determined the convergence rates for (1.1) and (1.2) for non-lattice

distributions  $F$  and for a wide but not exhaustive class of lattice distributions  $F$ . The objective of this note is to establish the convergence rates for the remaining lattice distributions  $F$ .

In stating our results we shall refer to the following list of possible assumptions about  $F$  and  $g$ :

*Assumption 1:* There exists  $\xi_0 \neq 0$  in  $I^0$  for which  $K'(\xi_0) = \epsilon$ .

*Assumption 1':* There exists  $\xi_1 \neq 0$  in  $I^0$  for which  $K(\xi_1) = \epsilon\xi_1$ .

*Assumption 2a:*  $F$  is non-lattice.

*Assumption 2b:*  $F$  is concentrated on a lattice  $\{b + jh : j \in \mathbb{Z}\}$  having span  $h > 0$   
and (for definiteness)  $b \in [0, h)$ .

*Assumption 2b':* Assumption 2b is met,  $\epsilon$  is in the lattice for  $F$ , and  $\alpha$  is an integer  
multiple of  $h$ .

From standard facts about the cgf

$$K_\epsilon(\xi) = K(\xi) - \epsilon\xi$$

for  $\mathcal{L}(X - \epsilon)$  it follows that 1' implies 1. that if 1 holds then  $\xi_0$  is unique and positive, and that if 1' holds then  $\xi_1 > \xi_0 > 0$  is also unique. One can show easily that if  $0 \in I^0$  then 1' holds for all small  $\epsilon > 0$ . Assumptions 2a and 2b are mutually exclusive and exhaustive possibilities, and 2b' is a special case of 2b.

Fill (1980, theorem 3.4.1; essentially repeated as Theorem 0 below) determined the convergence rate for (1.1) (respectively, (1.2)) when assumptions 0, 1 (respectively, 1'), and either 2a or 2b' are met. In Theorem 1 below we extend Fill's result under assumption 2b' to the more general case 2b. When  $\alpha = 0$  our result recaptures the principal result of Bahadur and Ranga Rao (1960) for (1.1) in their "case 2" (see also Blackwell and Hodges (1959)) and extends theorem 1 in Siegmund (1975) dealing with (1.2). When comparing (1.9) and (1.20) below to the results of Bahadur and Ranga Rao, note the strict inequality in our formulation (1.1) and the weak inequality in theirs; the difference is addressed in section 4.2 of Fill (1980) and in

Section 2.1 below. When comparing (1.9) and (1.21) below to Siegmund's theorem, it is important to note that (i) the factor  $(\exp(\theta_1 h) - 1)$  appearing on the right in his (14) should be  $(1 - \exp(-\theta_1 h))$ , and (ii) in his (11) Siegmund tacitly assumes (see the proof of his lemma 1) that  $h$  is the span of  $F$ . Consequently when  $\alpha = 0$  our assumption 2b' amounts to Siegmund's (11). For given  $F$  satisfying assumption 2b, assumption 2b' precludes all values of  $\epsilon$  smaller than  $b$ ; this is a severe restriction indeed since the most interesting values of  $\epsilon$  (in relation to the laws of large numbers) are values near zero.

Even when  $\alpha = 0$  our new result for (1.2) handles some important lattice distributions  $F$  for which Siegmund's result says nothing. For example in the case ( $X = \pm 1$  with probability 1/2 each) of simple symmetric random walk,  $K(\xi) = \log \cosh \xi$  for  $\xi \in \mathbb{R}$  and each of the assumptions 1 and 1' is equivalent to  $\epsilon < 1$ . Then assumption 2b' does not hold, but of course the more general 2b does, with  $h = 2$ .

In a forthcoming paper (Fill, 1987) one of the present authors will present a new argument due in part to David Siegmund yielding the convergence rate – and under further conditions an asymptotic expansion – for (1.2) in cases 2a and 2b when assumption 1' is weakened to assumption 1.

To state our results we need some additional notation. Suppose assumptions 0 and 1 are met. Let  $P_0$  denote the probability measure under which  $X, X_1, X_2, \dots$  are i.i.d. with

$$(1.3) \quad P_0\{X \in dx\} = \exp[\xi_0 x - K_\epsilon(\xi_0)]P\{X - \epsilon \in dx\}.$$

$P_0$  has cdf

$$(1.4) \quad \phi(\theta) = K_\epsilon(\xi_0 + \theta) - K_\epsilon(\xi_0) = K(\xi_0 + \theta) - K(\xi_0) - \epsilon\theta,$$

mean

$$(1.5) \quad \phi'(0) = K'_\epsilon(\xi_0) = 0,$$

and variance

$$(1.6) \quad \sigma^2 := \phi''(0) = K''_\epsilon(\xi_0) = K''(\xi_0).$$

Put

$$(1.7) \quad \theta_0 = -\xi_0 < 0;$$

then

$$(1.8) \quad \phi(\theta_0) = -K_\epsilon(\xi_0) = \frac{1}{2}\epsilon^2 - \epsilon^3\lambda(\epsilon) > 0,$$

where  $\lambda$  is the Cramér series for  $F$ .

**Theorem 0.** (Fill, 1980) If assumptions 0, 1, and either 2a or 2b' are met, then as  $m \rightarrow \infty$

$$(1.9) \quad P\{S_m > g(m)\} \sim a_m.$$

If assumption 1' also holds, then as  $m \rightarrow \infty$

$$(1.10) \quad p_m \sim (1 + \gamma)a_m.$$

Here

$$(1.11) \quad a_m := C(2\pi\sigma^2 m)^{-1/2} \exp(-m\phi(\theta_0) - |\theta_0|\alpha)$$

and

$$(1.12) \quad 0 < \gamma := \exp \left\{ \sum_{n=1}^{\infty} n^{-1} [\epsilon^{-n\phi(\theta_0)} P_0\{S_n > 0\} - P\{S_n > \epsilon n\}] \right\} - 1,$$

where if assumption 2a is met

$$(1.13a) \quad C = |\theta_0|^{-1}$$

while if assumption 2b' is met

$$(1.13b) \quad C = h[\exp(|\theta_0|h) - 1]^{-1}.$$

When assumption 2b holds we define for every  $m$

$$(1.14) \quad \delta_m = \text{fractional part of } h^{-1}(g(m) - bm).$$

$$(1.15) \quad \epsilon_m = \exp(|\theta_0| h \delta_m) - 1,$$

$$(1.16) \quad \lambda_m = (1 + \gamma)^{-1} \exp \left[ - \sum_{n=1}^{\infty} n^{-1} P\{S_n > \epsilon n\} \right] \sum_{n=0}^{\infty} \epsilon_{m+n} u_n = \sum_{n=0}^{\infty} \epsilon_{m+n} u_n / \sum_{n=0}^{\infty} u_n,$$

where  $\gamma$  is given by (1.12) and  $(u_n)$  has generating function

$$(1.17) \quad \tilde{u}(z) = \sum_{n=0}^{\infty} u_n z^n = \exp \left[ \sum_{k=1}^{\infty} z^k k^{-1} \exp(-k\phi(\theta_0)) P_0\{S_k > 0\} \right],$$

valid for all complex  $z$  with  $|z| \leq 1$ .

Notice that when assumption 2b' holds,  $\delta_m$ ,  $\epsilon_m$ , and  $\lambda_m$  all vanish identically. In general,  $0 \leq \delta_m < 1$ ,

$$(1.18) \quad 0 \leq \epsilon_m < \exp(|\theta_0| h) - 1,$$

and

$$(1.19) \quad 0 \leq \lambda_m < \exp(|\theta_0| h) - 1.$$

We now state the principal result of this note.

**Theorem 1.** If assumptions 0, 1, and 2b are met, then as  $m \rightarrow \infty$

$$(1.20) \quad P\{S_m > g(m)\} \sim a_m(1 + \epsilon_m).$$

If assumption 1' also holds, then as  $m \rightarrow \infty$

$$(1.21) \quad p_m \sim (1 + \gamma)a_m(1 + \lambda_m).$$

Here we have used the notation of (1.11), (1.12), (1.13b), (1.15), and (1.16).

**Remark.** How big are the quantities involved in Theorems 0 and 1 when  $\epsilon$  is small?

(a) One can show easily for any  $F$  satisfying assumption 0 and  $0 \in I^0$  that as  $\epsilon \downarrow 0$

$$(1.22) \quad |\theta_0| \sim \epsilon, \quad C \sim \epsilon^{-1}, \quad \sigma^2 \rightarrow 1, \quad \phi(\theta_0) \sim \frac{1}{2}\epsilon^2,$$

where  $C$  is given by (1.13a) or (1.13b) according as assumption 2a or 2b is met. As Siegmund (1975, remark 4(b)) notes,

$$(1.23) \quad \gamma \rightarrow 1 \text{ as } \epsilon \downarrow 0.$$

We identify  $\gamma$  up through order  $\epsilon$  in Section 5.1 below.

(b) In light of (1.18)–(1.19) we find

$$(1.24) \quad 0 \leq \epsilon_m = O(\epsilon) = o(1) \text{ uniformly in } m \text{ as } \epsilon \downarrow 0$$

and likewise

$$(1.25) \quad 0 \leq \lambda_m = O(\epsilon) = o(1) \text{ uniformly in } m \text{ as } \epsilon \downarrow 0$$

when  $F$  satisfies assumption 0 and  $0 \in I^0$ . We characterize the order- $\epsilon$  behavior of the sequences  $(\epsilon_m)$  and  $(\lambda_m)$  in Section 5.2 below. In particular we obtain the result

$$(1.26) \quad \frac{\lambda_m}{\epsilon h} \rightarrow \frac{1}{2} \text{ uniformly in } m \text{ as } \epsilon \downarrow 0,$$

showing that even on the scale of  $\epsilon$  the factor  $1 + \lambda_m$  added in going from (1.10) to (1.21) varies little as  $m$  changes when  $\epsilon$  is small. ■

The proof of Theorem 1 is given in Section 2. Sections 3 and 4 discuss further the situation in which  $h^{-1}(\epsilon - b)$  is irrational and rational, respectively. In particular, when  $h^{-1}(\epsilon - b)$  is rational we simplify the formulation (1.16)–(1.17) of  $\lambda_m$ : see (4.10)–(4.11). In Section 5 we analyze  $\gamma$ ,  $(\epsilon_m)$ , and  $(\lambda_m)$  for small  $\epsilon$ .

## 2. Proof of Theorem 1.

**2.0. Preliminaries.** To prepare for the proof of Theorem 1 we introduce the distributions associated through “exponential tilting” with that of  $X$  under  $P_0$ . This is accomplished by (8) in Siegmund (1975), but we describe the setup here for completeness. For each real  $\theta$  for which  $\phi(\theta) < \infty$ , let  $P_\theta$  denote the probability under which  $X, X_1, X_2, \dots$  are i.i.d. with

$$(2.1) \quad \begin{aligned} P_\theta\{X \in dx\} &= \exp[\theta x - \phi(\theta)]P_0\{X \in dx\} \\ &= \exp[\theta x - \phi(\theta)]\exp[\xi_0 x - K_\epsilon(\xi_0)]P\{X - \epsilon \in dx\}. \end{aligned}$$

(Since (see (1.4))  $\phi(0) = 0$ , the definition of  $P_0$  is consistent with (1.3).) The corresponding cgf  $\phi_\theta$  is given by

$$(2.2) \quad \begin{aligned} \phi_\theta(\eta) &= \phi(\theta + \eta) - \phi(\theta) = K_\epsilon(\xi_0 + \theta + \eta) - K_\epsilon(\xi_0 + \theta) \\ &= K(\xi_0 + \theta + \eta) - K(\xi_0 + \theta) - \epsilon\eta. \end{aligned}$$

In particular,

$$(2.3) \quad \begin{aligned} E_\theta X &= \phi'_\theta(0) = \phi'(\theta) &<, =, \text{ or } > 0 \\ \text{according as } \theta & &<, =, \text{ or } > 0 \end{aligned}$$

by (1.5) and the strict convexity of  $\phi$ .

We shall have particular interest in the values

$$(2.4) \quad \theta_0 = -\xi_0 < 0 \quad (\text{recall (1.7)}), \quad \theta_1 = \xi_1 - \xi_0 > 0$$

(the latter being defined only if assumption 1' is met), which satisfy

$$(2.5) \quad \phi(\theta_0) = \phi(\theta_1) = -K_\epsilon(\xi_0) = -(K(\xi_0) + \epsilon\theta_0).$$

As special cases of (2.1),

$P_0$  is given by (1.3),

$$P_{\theta_1}\{X \in dx\} = \exp(\xi_1 x)P\{X - \epsilon \in dx\}.$$

and

$$(2.6) \quad P_{\theta_0}\{X \in dx\} = P\{X - \epsilon \in dx\}.$$

Thus we can write our fundamental probabilities (1.1) and (1.2) in the forms

$$(2.7) \quad P\{S_m > g(m)\} = P_{\theta_0}\{S_m > \alpha\},$$

$$(2.8) \quad p_m = P_{\theta_0}\{S_n > \alpha \text{ for some } n \geq m\}.$$

Without loss of generality we take  $P_\theta$  itself to be the distribution of  $X_1, X_2, \dots$  defined on the space of (infinite) sequences of real numbers. Accordingly, let  $P_\theta^{(n)}$  denote the restriction of  $P_\theta$  to the  $\sigma$ -algebra generated by the first  $n$  coordinates ( $n = 1, 2, \dots$ ). Then for any  $\theta'$  and  $\theta''$ ,  $P_{\theta'}$  and  $P_{\theta''}$  are mutually absolutely continuous, and by (2.1)

$$(2.9) \quad \frac{dP_{\theta'}^{(n)}}{dP_{\theta''}^{(n)}} = \exp\{(\theta' - \theta'')S_n - n[\phi(\theta') - \phi(\theta'')]\}.$$

In particular, by (2.4), (2.5), and (2.9),

$$(2.10) \quad \frac{dP_{\theta_0}^{(n)}}{dP_{\theta_1}^{(n)}} = \exp(-\xi_1 S_n) \quad (n = 1, 2, \dots).$$

**2.1. Proof of (1.20).** The idea is to tilt from  $P_{\theta_0}$  in (2.7) to the distribution  $P_0$  under which  $X$  (and hence each  $S_m$ ) has mean zero (recall (2.3)) and then apply the (local) central limit theorem (CLT) for lattice distributions. According to (2.9) and (1.7)

$$(2.11) \quad P\{S_m > g(m)\} = P_{\theta_0}\{S_m > \alpha\} = \exp(-m\phi(\theta_0))E_0(\exp(-|\theta_0|S_m); S_m > \alpha).$$

The analysis of  $E_0(\exp(-|\theta_0|S_m); S_m > \alpha)$  can be carried out just as for  $E_0(\exp(\theta_1 S_m); S_m \leq hk)$  in Siegmund (1975, lemma 1(b)); the result is

$$(2.12) \quad E_0(\exp(-|\theta_0|S_m); S_m > \alpha) \sim C(2\pi\sigma^2 m)^{-1/2} \exp(-|\theta_0|\alpha)(1 + e_m).$$

Combination of (2.11) and (2.12) gives (1.20) immediately.

We remark that the corresponding result for the events  $\{S_m \geq g(m)\}$  is

$$(2.13) \quad P\{S_m \geq g(m)\} \sim a_m(1 + \tilde{\epsilon}_m)$$

where

$$(2.14) \quad \tilde{\delta}_m = 1 - \text{fractional part of } h^{-1}(bm - g(m)),$$

$$(2.15) \quad \tilde{\epsilon}_m = \exp(|\theta_0|h\tilde{\delta}_m) - 1.$$

The probabilities  $P\{S_m > g(m)\}$  and  $P\{S_m \geq g(m)\}$ , and likewise the right sides of (1.20) and (2.13), differ only when  $g(m)$  is a point in the lattice for  $S_m$ . In that case the right side of (2.13) is larger than the right side of (1.20) by the factor  $\exp(|\theta_0|h)$ . This remark generalizes (4.2.13) in Fill (1980).

## 2.2. Proof of (1.21). Let

$$(2.16) \quad T_m = \inf\{n : n \geq m, S_n > \alpha\},$$

the inf of the empty set being  $+\infty$ . Then by (2.8)

$$(2.17) \quad p_m = P_{\theta_0}\{S_m > \alpha\} + P_{\theta_0}\{m < T_m < \infty\} = P\{S_m > g(m)\} + P_{\theta_0}\{m < T_m < \infty\}.$$

The first term was analyzed in Section 2.1; the result is (1.20). For the second term we use the approach of Siegmund (1975) exploited also by Fill (1980, 1983):

- (a) Tilt from  $P_{\theta_0}$  to  $P_{\theta_1}$  using the fundamental identity of sequential analysis (FISA), and simplify using the SLLN.
- (b) Apply the weak-law result (2.13) to estimate  $P_{\theta_1}\{S_m \leq \alpha\}$  and to approximate the conditional distribution under  $P_{\theta_1}$  of  $S_m$  given  $S_m \leq \alpha$ .
- (c) Complete the proof via renewal-theoretic calculations.

The first two steps are routine; the last involves a considerable complication of Siegmund's arguments and is the object of Section 2.3 below.

Step (a) is carried out just as for (22)–(23) in Siegmund (1975). The result is

$$\begin{aligned}
 P_{\theta_0}\{m < T_m < \infty\} &= \exp(-\xi_1 \alpha) E_{\theta_1}(\exp[-\xi_1(S_{T_m} - \alpha)]; m < T_m < \infty) \\
 &= \exp(-\xi_1 \alpha) E_{\theta_1}(\exp[-\xi_1(S_{T_m} - \alpha)]; S_m \leq \alpha) \\
 &= \exp(-\xi_1 \alpha) P_{\theta_1}\{S_m \leq \alpha\} \\
 &\quad \times \int_{[0, \infty)} E_{\theta_1}(\exp[-\xi_1(S_{T_m} - \alpha)] \mid S_m = \alpha - y) P_{\theta_1}\{S_m \in \alpha - dy \mid S_m \leq \alpha\} \\
 (2.18) \quad &= \exp(-\xi_1 \alpha) P_{\theta_1}\{S_m \leq \alpha\} \\
 &\quad \times \int_{[0, \infty)} E_{\theta_1} \exp[-\xi_1(S_{\tau(y)} - y)] P_{\theta_1}\{\alpha - S_m \in dy \mid S_m \leq \alpha\}
 \end{aligned}$$

where, for  $y \geq 0$ ,  $\tau(y) = \inf\{n : S_n > y\}$ .

Next we perform step (b). The result (2.13) can be applied to

$$P_{\theta_1}\{S_m \leq \alpha\} = P_{\theta_1}\left\{\sum_{k=1}^m [\sigma_{\theta_1}^{-1}(-X_k + \mu_{\theta_1})] \geq -\sigma_{\theta_1}^{-1}\alpha + \sigma_{\theta_1}^{-1}\mu_{\theta_1}m\right\}$$

directly, where

$$\mu_{\theta_1} = E_{\theta_1} X > 0 \quad (\text{recall (2.3)–(2.4)}), \quad \sigma_{\theta_1}^2 = \text{Var}_{\theta_1} X > 0.$$

However, it is probably easier to start afresh and perform calculations (tilt from  $P_{\theta_1}$  to  $P_0$ , apply the CLT) parallel to those for (2.13) or (1.20). Either way one finds, making use of (2.5).

$$(2.19) \quad P_{\theta_1}\{S_m \leq \alpha\} \sim [\exp(|\theta_0|h) - 1][1 - \exp(-\theta_1 h)]^{-1} \exp[(\theta_1 + |\theta_0|)\alpha] a_m \exp(-\theta_1 h \delta_m).$$

Similarly, for each fixed  $j \in \mathbb{Z}$

$$(2.20) \quad P_{\theta_1}\{\alpha - S_m \geq h(j + \delta_m)\} \sim \exp(-\theta_1 h j) \times \quad (\text{right side of (2.19)}).$$

Dividing (2.20) by (2.19) we obtain the analogue

$$(2.21) \quad \lim_{m \rightarrow \infty} P_{\theta_1}\{\alpha - S_m \geq h(j + \delta_m) \mid S_m \leq \alpha\} = \exp(-\theta_1 h j) \quad (j = 0, 1, \dots)$$

to (38) in Siegmund (1975). In words, (2.21) says that under  $P_{\theta_1}$  the conditional distribution of  $\alpha - S_m - h\delta_m$  given  $S_m \leq \alpha$  converges to the geometric distribution with probability mass

function given by

$$[1 - \exp(-\theta_1 h)] \exp(-\theta_1 h j) \quad (j = 0, 1, \dots).$$

It follows that the last integral appearing in (2.18) equals

$$(2.22) \quad [1 - \exp(-\theta_1 h)] \exp(\theta_1 h \delta_m) Z_m^* + o(1),$$

where we define for  $x \geq 0$

$$(2.23) \quad Z(x) = E_{\theta_1} \exp[-\xi_1(S_{\tau(x)} - x)]$$

and for each  $n$

$$(2.24) \quad B_n = \{h(j + \delta_n) : j \in \mathbb{Z}\} = \text{(lattice for } \mathcal{L}_{\theta_1}(\alpha - S_n)),$$

$$(2.25) \quad Z_n^* = \sum_{0 \leq x \in B_n} Z(x) \exp(-\theta_1 x).$$

Notice that when assumption 2b' is met,  $B_n \equiv B$  consists of the multiples of  $h$  and  $Z_n^* \equiv Z^*$  reduces to (the lattice version of)  $Z^*(\theta_1)$  in Siegmund (1975). In Section 2.3 below we show

$$(2.26) \quad Z_m^* = [\exp(|\theta_0| h) - 1]^{-1} [\gamma + (1 + \gamma)\lambda_m - \epsilon_m].$$

Combining (2.18), (2.19), (2.22) and (2.26) we then find

$$(2.27) \quad P_{\theta_0} \{m < T_m < \infty\} = a_m [\gamma + (1 + \gamma)\lambda_m - \epsilon_m + o(1)].$$

Thus by (2.17) and (1.20)

$$(2.28) \quad p_m = a_m [(1 + \gamma)(1 + \lambda_m) + o(1)] = (1 + o(1))(1 + \gamma)a_m(1 + \lambda_m),$$

which is (1.21).

Before examining the proof of (2.26) we close Section 2.2 with two remarks concerning (1.21).

**Remark.** (a) One can show that for fixed  $\alpha \in \mathbb{R}$  the last integral appearing in (2.18) equals (2.22) uniformly for  $\epsilon \in (0, \epsilon_1]$ , where  $\epsilon_1$  is any  $\epsilon > 0$  for which assumption 1' is met. This generalizes a remark of Siegmund (1975).

(b) (2.13) has the strong-law counterpart

$$(2.29) \quad p_m^+ := P\{S_n \geq g(n) \text{ for some } n \geq m\} \sim (1 + \gamma)a_m(1 + \bar{\lambda}_m),$$

where  $\bar{\lambda}_m$  is defined by the right side of (1.16) when  $e_{m+n}$  there is replaced by  $\tilde{e}_{m+n}$  of (2.14)–(2.15).  $\bar{\lambda}_m$  is bigger than  $\lambda_m$  by the amount  $(1 + \gamma)^{-1} \exp[-\sum_{n=1}^{\infty} n^{-1} P\{S_n > \epsilon n\}] [\exp(|\theta_0| h) - 1]$  times

$$(2.30) \quad \sum_{n \geq 0: \delta_{m+n}=0} u_n.$$

When assumption 2b' is met, (2.30) equals

$$(2.31) \quad \begin{aligned} \sum_{n=0}^{\infty} u_n &= \bar{u}(1) = \exp \left[ \sum_{n=1}^{\infty} n^{-1} \exp(-n\phi(\theta_0)) P_0\{S_n > 0\} \right] \\ &= (1 + \gamma) \exp \left[ \sum_{n=1}^{\infty} n^{-1} P\{S_n > \epsilon n\} \right] \end{aligned}$$

so that

$$(2.32) \quad p_m^+ \sim \exp(|\theta_0| h) p_m,$$

which is (4.2.14) in Fill (1980). At the other extreme, when, for example,  $\alpha = 0$  and  $h^{-1}(\epsilon - b)$  is irrational, (2.30) vanishes. Indeed, we have in that case  $p_m^+ = p_m$  for every  $m$ . ■

**2.3. Proof of (2.26): renewal-theoretic calculations.** Define  $G$  to be the distribution of  $S_{\tau(0)}$  under  $P_{\theta_1}$ , with Laplace transform

$$(2.33) \quad G^*(\lambda) := E_{\theta_1} \exp(-\lambda S_{\tau(0)}) = \int_{(0, \infty)} \exp(-\lambda x) G(dx).$$

For each  $n$  define

$$(2.34) \quad G_n(y) = P_{\theta_1}\{\tau(0) = n, S_n \leq y\},$$

$$(2.35) \quad G_n^*(\lambda) = E_{\theta_1}(\exp(-\lambda S_{\tau(0)}); \tau(0) = n) = \int_{(0,\infty)} \exp(-\lambda x) G_n(dx).$$

As Siegmund shows (when  $\alpha = 0$ )

$$(2.36) \quad Z^* = |\theta_0|^{-1}[G^*(\theta_1) - G^*(\xi_1)] + Z^* G^*(\theta_1)$$

for  $Z^* := \int_{(0,\infty)} Z(x) \exp(-\theta_1 x) dx$  in case 2a and (tacitly)

$$(2.37) \quad Z^* = [\exp(|\theta_0|h) - 1]^{-1}[G^*(\theta_1) - G^*(\xi_1)] + Z^* G^*(\theta_1)$$

for  $Z^*$  (like  $B_n$ , independent of  $n$ ) given by (2.25) in case 2b', so we shall derive

$$(2.38) \quad Z_m^* = z_m + \sum_{n=0}^{\infty} Z_{m+n}^* g_n, \quad m \geq 1.$$

in the present situation, where for each  $m \geq 0$

$$(2.39) \quad g_m := G_m^*(\theta_1) \text{ (in particular, } g_0 = 0\text{),}$$

$$(2.40) \quad z_m := [\exp(|\theta_0|h) - 1]^{-1}[G^*(\theta_1) - G^*(\xi_1) + \sum_{n=1}^{\infty} e_{m+n} g_n - e_m G^*(\xi_1)].$$

Here is the argument leading to (2.38). Begin with the following renewal equation for the function  $Z$  of (2.23):

$$\begin{aligned} Z(x) &= \int_{(x,\infty)} \exp[-\xi_1(y-x)] G(dy) + \int_{(0,x]} Z(x-y) G(dy) \\ &= \sum_{n=0}^{\infty} \left[ \sum_{y:y>x, \alpha-y \in B_n} \exp[-\xi_1(y-x)] P_{\theta_1}\{S_{\tau(0)} = y; \tau(0) = n\} \right. \\ &\quad \left. + \sum_{y:0<y \leq x, \alpha-y \in B_n} Z(x-y) P_{\theta_1}\{S_{\tau(0)} = y; \tau(0) = n\} \right]; \end{aligned}$$

compare (26) in Siegmund (1975). Multiply both sides by  $\exp(-\theta_1 x)$ , sum over values  $x$  satisfying  $0 \leq x \in B_m$ , and apply to the second sum on  $y$  the observation that when  $\alpha - y \in B_n$ ,

$x \in B_m$  if and only if  $x - y \in B_{m+n}$  to conclude

$$Z_m^* = \sum_{n=0}^{\infty} \left[ \sum_{y:y>0, \alpha-y \in B_n} \exp(-\xi_1 y) P_{\theta_1} \{S_{\tau(0)} = y; \tau(0) = n\} \right. \\ \times \left. \sum_{x:0 \leq x < y, x \in B_m} \exp(|\theta_0|x) + Z_{m+n}^* g_n \right].$$

Now the sum on  $x$  here equals  $[\exp(|\theta_0|h) - 1]^{-1}(1 + e_m)$  times

$$\exp(|\theta_0|h[h^{-1}y - \delta_m]) - 1 = (1 + e_m)^{-1} \exp(|\theta_0|y) \exp[|\theta_0|h(h^{-1}y - \delta_m) - (h^{-1}y - \delta_m)] - 1,$$

and  $[h^{-1}y - \delta_m] - (h^{-1}y - \delta_m)$  is for  $\alpha - y \in B_n$  the fractional part of each of the numbers  $\delta_m - h^{-1}y = \delta_m - h^{-1}\alpha + h^{-1}(\alpha - y)$ ,  $\delta_m + \delta_n - h^{-1}\alpha$ , and  $h^{-1}(g(m+n) - b(m+n))$  – namely,  $\delta_{m+n}$ . Thus

$$\sum_{x:0 \leq x < y, x \in B_m} \exp(|\theta_0|x) = [\exp(|\theta_0|h) - 1]^{-1}[(1 + e_{m+n}) \exp(|\theta_0|y) - (1 + e_m)],$$

and (2.38) now follows directly.

Whereas solving (2.36) or (2.37) for  $Z^*$  is a trivial matter, solving (2.38) requires some work. We observe first that the sequence  $(g_m)_{m \geq 0}$  is a sub-probability mass function with total mass  $G^*(\theta_1) < 1$ , and that the sequences  $(z_m)$  and  $(Z_m^*)$  are bounded. Let  $u$  be the renewal mass function associated with  $g$ , to wit:

$$(2.41) \quad u_m = \sum_{j=0}^{\infty} g_m^{(j)},$$

where  $g^{(j)}$  is the  $j$ th convolution power of  $g$ ,  $g^{(0)}$  placing unit mass at 0. We show below that  $(u_m)$  has generating function  $\hat{u}$  given by (1.17). Applying the method of successive substitutions to (2.38) in the same way that one discovers the solution to a renewal equation, one finds that

$$(2.42) \quad Z_m^* = \sum_{n=0}^{\infty} z_{m+n} u_n, \quad m \geq 1,$$

the series converging absolutely, is the unique bounded solution to (2.38).

We shall reduce (2.42) to (2.26), but first we calculate the generating functions for the

sequences  $g$  and  $u$ . The generating function for  $g$  is

$$(2.43) \quad \begin{aligned} \tilde{g}(w) &= \sum_{n=0}^{\infty} g_n w^n = \sum_{n=1}^{\infty} E_{\theta_1}(\exp(-\theta_1 S_{\tau(0)}); \tau(0) = n) w^n \\ &= E_{\theta_1}[w^{\tau(0)} \exp(-\theta_1 S_{\tau(0)})]. \end{aligned}$$

According to a result of Spitzer (compare Chung (1974, theorem 8.4.2))

$$(2.44) \quad 1 - E_{\theta_1}[w^{\tau(0)} \exp(-\lambda S_{\tau(0)})] = \exp \left[ - \sum_{k=1}^{\infty} w^k k^{-1} E_{\theta_1}(\epsilon^{-\lambda S_k}; S_k > 0) \right].$$

Equations (2.43) and (2.44) are valid for all complex  $w$  with  $|w| \leq 1$ . From (2.9) and (2.5) follows

$$(2.45) \quad E_{\theta_1}(\epsilon^{-\theta_1 S_k}; S_k > 0) = \exp(-k\phi(\theta_0)) P_0\{S_k > 0\}.$$

Combining (2.43)–(2.45).

$$(2.46) \quad \tilde{g}(w) = 1 - \exp \left[ - \sum_{k=1}^{\infty} w^k k^{-1} \exp(-k\phi(\theta_0)) P_0\{S_k > 0\} \right], \quad |w| \leq 1.$$

Now by (2.41),  $\tilde{u}(w) = (1 - \tilde{g}(w))^{-1}$ , and (1.17) is established. In particular, (2.31) holds.

Combining (2.42), (2.40), (2.31), and the consequences

$$(2.47) \quad 1 - G^*(\theta_1) = \exp \left[ - \sum_{k=1}^{\infty} k^{-1} \exp(-k\phi(\theta_0)) P_0\{S_k > 0\} \right],$$

$$(2.48) \quad 1 - G^*(\xi_1) = \exp \left[ - \sum_{k=1}^{\infty} k^{-1} P\{S_k > \epsilon k\} \right]$$

of (2.44), (2.45), and (2.6), we find

$$(2.49) \quad Z_m^* = [\exp(|\theta_0| h) - 1]^{-1} \left[ \gamma + \sum_{n=1}^{\infty} v_{m+n} g_n - v_m G^*(\xi_1) \right]$$

where

$$(2.50) \quad v_m := \sum_{n=0}^{\infty} \epsilon_{m+n} u_n.$$

But by (2.41)

$$(2.51) \quad \sum_{n=1}^{\infty} v_{m+n} g_n = \sum_{n=0}^{\infty} e_{m+n} (u * g)_n = \sum_{n=1}^{\infty} e_{m+n} u_n = v_m - e_m,$$

so by (2.48)–(2.51) and the definition (1.16) of  $\lambda_m$ , (2.26) holds.

### 3. Theorem 1 when $r = h^{-1}(\epsilon - b)$ is irrational.

In this section we examine Theorem 1 under the assumption that  $r = h^{-1}(\epsilon - b)$  is irrational. Specifically, we first recall that the sequences  $(\epsilon_m)$  and  $(\lambda_m)$  appearing in (1.20) and (1.21) take values in the interval  $[0, \exp(|\theta_0|h) - 1]$ . We shall now say how the values of each sequence are distributed over this interval. We shall use this information to compute the respective limits as  $M \rightarrow \infty$  of the average values  $M^{-1} \sum_{m=1}^M \epsilon_m$  and  $M^{-1} \sum_{m=1}^M \lambda_m$ .

Recall that a sequence  $(x_m)$  of real numbers is said to have distribution  $F$  if, letting  $m(M)$  be a random variable uniformly distributed on  $\{1, \dots, M\}$ , the random variables  $x_{m(M)}$  converge in distribution to  $F$  as  $M \rightarrow \infty$ . If, in particular,  $F$  is the uniform distribution on an interval  $I$ ,  $(x_m)$  is said to be equidistributed over  $I$ .

According to a well-known result, the sequence  $\delta_m = [\text{fractional part of } (h^{-1}\alpha + rm)]$  of (1.14) is equidistributed on  $[0, 1)$ . Thus the distribution of the sequence  $(\epsilon_m)$  of (1.15) is that of

$$(3.1) \quad V = \exp(|\theta_0|hU) - 1,$$

where  $U$  is uniformly distributed on  $[0, 1)$ . In particular,

$$(3.2) \quad \bar{\epsilon}_M := M^{-1} \sum_{m=1}^M \epsilon_m - E[\exp(|\theta_0|hU) - 1] = (|\theta_0|h)^{-1}[\exp(|\theta_0|h) - 1] - 1 = (|\theta_0|C)^{-1} - 1$$

with  $C$  given by (1.13b). It is interesting to note that the corresponding average  $C(1 + \bar{\epsilon}_M)$  of the coefficients  $C(1 + \epsilon_m)$  multiplying the expression  $(2\pi\sigma^2m)^{-1/2} \exp(-m\phi(\theta_0) - |\theta_0|\alpha)$  in (1.20) tends as  $M \rightarrow \infty$  to the constant coefficient  $|\theta_0|^{-1}$  multiplying the same expression in (1.9) in the non-lattice case.

We next characterize the distribution of  $(\lambda_m)$ . We begin by recalling from (1.16) that  $\lambda_m$  can be written in the form

$$(3.3) \quad \lambda_m = \sum_{n=0}^{\infty} \epsilon_{m+n} u_n^0 = E \epsilon_{m+W} = \sum_{n=0}^{\infty} [\exp(|\theta_0|h(\delta_m + nr)) - 1] u_n^0,$$

where  $W$  is a random variable with probability mass function (pmf)

$$(3.4) \quad u_n^0 := u_n / \tilde{u}(1), \quad n = 0, 1, \dots,$$

and where  $\langle x \rangle$  for real  $x$  denotes the fractional part of  $x$ . Since the mapping  $t : [0, 1) \rightarrow [0, 1)^\infty$  defined by

$$t(u) = (u, \langle u + r \rangle, \langle u + 2r \rangle, \dots) \quad (u \in [0, 1))$$

is continuous except at a countable number of points, it follows easily that the distribution of  $(\lambda_m)$  is that of

$$(3.5) \quad V^* = \sum_{n=0}^{\infty} [\exp(|\theta_0| h \langle U + nr \rangle) - 1] u_n^0 = E[\exp(|\theta_0| h \langle U + W r \rangle) | U] - 1,$$

where  $U \sim \text{unif}[0, 1)$  and  $W$  with pmf  $(u_n^0)_{n \geq 0}$  are independent random variables. In particular,

$$(3.6) \quad \bar{\lambda}_M := M^{-1} \sum_{m=1}^M \lambda_m = E \exp(|\theta_0| h \langle U + W r \rangle) - 1 = E \exp(|\theta_0| h U) - 1 = (|\theta_0| C)^{-1} - 1.$$

In (3.6) we have used the well-known fact that  $\langle U + Y \rangle \sim \text{unif}[0, 1)$  when  $U \sim \text{unif}[0, 1)$  and  $Y$  are independent. The average  $C(1 + \bar{\lambda}_M)$  of the coefficients  $C(1 + \lambda_m)$  multiplying the expression  $(1 + \gamma)(2\pi\sigma^2 m)^{-1/2} \exp(-m\phi(\theta_0) - |\theta_0|\alpha)$  in (1.21) tends as  $M \rightarrow \infty$  to the constant coefficient  $|\theta_0|^{-1}$  of the same expression in (1.10) in the non-lattice case.

#### 4. Theorem 1 when $r = h^{-1}(\epsilon - b)$ is rational.

In this section we examine Theorem 1 under the assumption that  $r = h^{-1}(\epsilon - b)$  is rational. The main achievement here is to express  $\lambda_m$  of (1.16) explicitly, rather than implicitly through the generating function  $\tilde{u}$  of (1.17). We also show that each of the sequences  $(e_m)$  and  $(\lambda_m)$  cycles through only  $q$  distinct values, where

$$(4.1) \quad r = h^{-1}(\epsilon - b) = p/q \text{ in lowest terms } (p \in \mathbb{Z}, q \in \mathbb{Z}),$$

and we show that the average of the  $q$  values is in either case

$$(4.2) \quad \bar{\epsilon} = \bar{\lambda} = \exp(|\theta_0|hq^{-1}\zeta) \times q^{-1}[\exp(|\theta_0|hq^{-1}) - 1]^{-1}[\exp(|\theta_0|h) - 1] - 1$$

with

$$(4.3) \quad \zeta := \text{fractional part of } [q \times (\text{fractional part of } h^{-1}\alpha)].$$

According to (1.16),  $(\lambda_m)$  is a constant multiple of the more simply defined sequence  $(v_m)$  of (2.50), so we shall work with the latter. Clearly each of the sequences  $(\delta_m)$ ,  $(e_m)$ ,  $(\lambda_m)$ , and  $(v_m)$  is periodic with period  $q$ . We shall write

$$(4.4) \quad \omega = e^{2\pi i/q}$$

and

$$(4.5) \quad \hat{x}_k = q^{-1} \sum_{j=0}^{q-1} x_j \omega^{-jk}$$

for the discrete Fourier transform of a periodic sequence  $(x_j)$  with period  $q$ . The inverse transform to (4.5) is

$$(4.6) \quad x_j = \sum_{k=0}^{q-1} \hat{x}_k \omega^{jk}.$$

We can write

$$(4.7) \quad v_m = \sum_{\ell=0}^{q-1} e_{m-\ell} w_\ell.$$

where

$$w_\ell := \sum_{0 \leq n \equiv -\ell \pmod{q}} u_n.$$

That is,  $v$  is the convolution of the periodic sequences  $e$  and  $w$ . Now the sequence  $w$  has Fourier transform

$$\hat{w}_k = q^{-1} \sum_{n=0}^{\infty} u_n \omega^{kn} = q^{-1} \hat{u}(\omega^k),$$

so

$$\hat{v}_k = q \hat{e}_k \hat{w}_k = \hat{e}_k \hat{u}(\omega^k)$$

and therefore

$$(4.8) \quad v_j = \sum_{k=0}^{q-1} \hat{e}_k \hat{u}(\omega^k) \omega^{jk}.$$

The factor  $\hat{u}(\omega^k)$  appearing in (4.8) can be explicitly presented. According to (1.17),

$$(4.9) \quad \hat{u}(\omega^k) = \exp \left[ \sum_{n=1}^{\infty} \omega^{kn} n^{-1} \exp(-n\phi(\theta_0)) P_0\{S_n > 0\} \right].$$

In summary,

$$(4.10) \quad \lambda_m = (1 + \gamma)^{-1} \exp \left[ - \sum_{n=1}^{\infty} n^{-1} P\{S_n > \epsilon n\} \right] \\ \times \sum_{k=0}^{q-1} \hat{e}_k \epsilon^{2\pi i km/q} \exp \left[ \sum_{n=1}^{\infty} e^{2\pi i kn/q} n^{-1} \exp(-n\phi(\theta_0)) P_0\{S_n > 0\} \right],$$

where

$$(4.11) \quad \hat{e}_k = q^{-1} \sum_{j=0}^{q-1} e_j e^{-2\pi i j k / q}.$$

In case 2b',  $q = 1$ ,  $\hat{e}_k \equiv 0$ , and  $\lambda_m \equiv 0$  (as already noted prior to (1.18)). The factor (4.11) can be simplified in certain other special subcases of assumption 2b, for example,  $\alpha = 0$  and  $p = 1$ . Indeed, we have then by (1.14)  $h^{-1}(g(m) - bm) = rm = m/q$ , and so  $\delta_m = q^{-1} \times (m \bmod q)$  and

$$\begin{aligned} \hat{e}_k &= q^{-1} \sum_{j=0}^{q-1} [\exp(|\theta_0| h q^{-1} j) - 1] \omega^{-jk} \\ &= q^{-1} [\exp(|\theta_0| h) - 1] [\exp(|\theta_0| h q^{-1}) \omega^{-k} - 1]^{-1} - \delta_{0,k}, \end{aligned}$$

where here  $\delta_{0,k} = 0$  or 1 according as  $k \neq 0$  or  $k = 0$ .

We turn our attention to the distribution of values of  $(e_m)$  and of  $(\lambda_m)$ . As  $m$  increases,  $e_m$  cycles through the values

$$(4.12) \quad \exp[|\theta_0| h \times q^{-1}(\zeta + j)] - 1, \quad j = 0, \dots, q - 1,$$

though not necessarily in this order. The average of these  $q$  values is given by (4.2). The  $q$  periodically repeating values of the sequence  $(\lambda_m)$  are presented explicitly in (4.10)–(4.11).

The average of any  $q$  consecutive values of  $v$  is

$$\hat{v}_0 = \hat{\epsilon}_0 \tilde{u}(1) = \bar{\epsilon}(1 + \gamma) \exp \left[ \sum_{n=1}^{\infty} n^{-1} P\{S_n > \epsilon n\} \right],$$

so the corresponding average for  $\lambda$  is  $\bar{\epsilon}$ .

## 5. Small- $\epsilon$ analysis.

Leaving aside the issue of how in general to compute the exact probabilities  $P\{S_n > \epsilon n\}$  and  $P_0\{S_n > 0\}$ , when  $\epsilon$  is small the constant  $\gamma = \gamma(\epsilon)$  of (1.12) that appears in (1.10) and (1.21) is difficult to compute because of the quite slow convergence of the series in (1.12). In Section 5.1 we provide an approximation to  $\gamma$  valid for any distribution  $F$ , non-lattice or lattice, when  $\epsilon$  is small. More precisely, assuming only that assumption 0 is met and  $0 \in I^0$ , we show that

$$(5.1) \quad \gamma(\epsilon) = 1 - (\rho(\epsilon) - \kappa/3)\epsilon + O(\epsilon^2),$$

where

$$(5.2) \quad \kappa := EX^3$$

is the third cumulant of  $F$  and

$$(5.3) \quad \rho(\epsilon) := E_{\theta_1} S_{\tau(0)}^2 / E_{\theta_1} S_{\tau(0)},$$

and that

$$(5.4) \quad \rho(\epsilon) \rightarrow \rho(0) := ES_{\tau(0)}^2 / ES_{\tau(0)}.$$

Thus

$$(5.5) \quad \gamma(\epsilon) = 1 - (\rho(0) - \kappa/3)\epsilon + o(\epsilon).$$

Spitzer (1960) and Lai (1976) have shown how  $ES_{\tau(0)}$  and  $ES_{\tau(0)}^2$  can be computed; see equations (1.2) and (1.5) in Lai (1976). Consequently,

$$(5.6) \quad \rho(0) = 2^{1/2}\mu_2 + \kappa/3 - 2 \sum_{n=1}^{\infty} n^{-1/2} [E(n^{-1/2}S_n)^- - (2\pi)^{-1/2}],$$

where

$$(5.7) \quad \mu_2 := 1 - \sum_{n=1}^{\infty} \left[ (\pi n)^{-1/2} - \binom{2n}{n} 2^{-2n} \right] = 0.8239168$$

is the value of  $ES_{\tau(0)}^2$  (and  $2^{1/2}\mu_2 = 1.1651943$  is the value of  $\rho(0)$ ) when  $F$  is the standard normal distribution function. Lai (1976, theorem 1) has shown that the series in (5.6) in fact converges absolutely. An alternative formula for  $\rho(0)$ , better suited for numerical calculations, has been given by Siegmund (1985, theorem 10.55 and problem 10.7):

$$(5.8) \quad \rho(0) = \kappa/3 - 2\pi^{-1} \int_0^\infty \lambda^{-2} \operatorname{Re} \log\{2[1 - f(\lambda)]/\lambda^2\} d\lambda,$$

where  $f$  is the characteristic function corresponding to  $F$ .

We return now to the lattice-case setting of Theorem 1. Exact computation of the  $\lambda_m$ 's of (1.16) is even more difficult than that of  $\gamma$ . Even in the relatively simple situation that  $r = h^{-1}(\epsilon - b)$  is rational, calculation of the sequence  $(\lambda_m)$  requires the evaluation of each of the  $q$  slowly converging (when  $\epsilon$  is small) series

$$(5.9) \quad \sum_{n=1}^{\infty} \epsilon^{2\pi i kn/q} n^{-1} \exp(-n\phi(\theta_0)) P_0\{S_n > 0\}, \quad k = 0, 1, \dots, q-1.$$

The main result of Section 5.2 establishes the strikingly simple approximation  $\lambda_m \doteq \epsilon h/2$  for all  $m$  made precise by (1.26) above.

**5.1. Small- $\epsilon$  analysis of  $\gamma$ .** Throughout Section 5.1 we suppose that assumption 0 is met and that  $0 \in I^0$ , so that assumptions 1 and 1' are met for all small  $\epsilon$ . One can easily establish

$$(5.10) \quad \xi_0 = \epsilon \left[ 1 - \frac{1}{2}\kappa\epsilon + O(\epsilon^2) \right], \quad \xi_1 = 2\epsilon \left[ 1 - \frac{2}{3}\kappa\epsilon + O(\epsilon^2) \right].$$

We begin the proof of (5.1) by recalling from (1.12), (2.33), and (2.47)–(2.48) that

$$(5.11) \quad \gamma = [1 - E_{\theta_1} \exp(-\xi_1 S_{\tau(0)})]/[1 - E_{\theta_1} \exp(-\theta_1 S_{\tau(0)})] - 1.$$

It then follows from Lemma 1 below and (5.10) and the definition (2.4) of  $\theta_1$  that

$$\begin{aligned} \gamma &= \left[ \xi_1 E_{\theta_1} S_{\tau(0)} - \frac{1}{2} \xi_1^2 E_{\theta_1} S_{\tau(0)}^2 + O(\epsilon^3) \right] / \left[ \theta_1 E_{\theta_1} S_{\tau(0)} - \frac{1}{2} \theta_1^2 E_{\theta_1} S_{\tau(0)}^2 + O(\epsilon^3) \right] - 1 \\ &= \left[ \xi_1 - \frac{1}{2} \xi_1^2 \rho(\epsilon) + O(\epsilon^3) \right] / \left[ \theta_1 - \frac{1}{2} \theta_1^2 \rho(\epsilon) + O(\epsilon^3) \right] - 1 \end{aligned}$$

$$(5.12) \quad = \left[ (\xi_1 - \theta_1) - \frac{1}{2}(\xi_1^2 - \theta_1^2)\rho(\epsilon) + O(\epsilon^3) \right] / \left[ \theta_1 - \frac{1}{2}\theta_1^2\rho(\epsilon) + O(\epsilon^3) \right] \\ = 1 - (\rho(\epsilon) - \kappa/3)\epsilon + O(\epsilon^2),$$

and (5.1) is established. The result (5.4) also follows from Lemma 1.

We state without proof a further asymptotic development of  $\gamma$  for two specific choices of  $F$ . When  $F$  is the standard normal distribution function, the remainder term in (5.5) equals  $(1 + o(1))(1 + \rho^2(0))\epsilon^2/4$  (recall  $\rho(0) = 1.1651943$  in this case). When  $F$  puts mass  $1/2$  each at  $\pm 1$ , so that  $S$  is simple symmetric random walk, we have

$$\gamma(\epsilon) = 1 - \epsilon + (1 + o(1))c\epsilon^{3/2},$$

where

$$c = \frac{2}{3} \left( \frac{2}{\pi} \right)^{1/2} \sum_{j=0}^{\infty} (4[(j+1)^{3/2} - j^{3/2}] - 3[(j+1)^{1/2} + j^{1/2}]) = 0.6634768.$$

For the cases of  $X \sim$  standard normal and  $P\{X = \pm 1\} = 1/2$  and selected values of  $\epsilon$ , the following table lists  $\gamma = \gamma(\epsilon)$  along with the errors in the corresponding approximations  $\hat{\gamma}(\epsilon) = 1 - \rho(0)\epsilon + (1 + \rho^2(0))\epsilon^2/4$  (normal case) and  $\hat{\gamma}(\epsilon) = 1 - \epsilon + c\epsilon^{3/2}$  (Bernoulli case).

$\epsilon$	$X \sim$ standard normal			$P\{X = \pm 1\} = 1/2$		
	$\gamma$	$\Delta = \hat{\gamma} - \gamma$	$\Delta/\epsilon^3$	$\gamma$	$\Delta = \hat{\gamma} - \gamma$	$\Delta/\epsilon^2$
.300	.6988447	.0046448	.172	.7579655	.0510548	.567
.200	.7891308	.0014071	.176	.8345596	.0247836	.620
.100	.8891949	.0001799	.180	.9148624	.0061186	.612
.050	.9431911	.0000227	.182	.9558413	.0015766	.631
.025	.9712357	.0000029	.183	.9772209	.0004017	.643
.010	.9884068	.0000002	.184	.9905982	.0000652	.652

In Lemmas 1 and 2 below we write  $\tau$  for the first ladder epoch  $\tau(0)$ .

**Lemma 1.** For any real  $r \geq 0$

$$(5.13) \quad E_{\theta_1} S_\tau^r \rightarrow E S_\tau^r \text{ as } \epsilon \downarrow 0.$$

*Proof.* We first relate the left side of (5.13) to the original random walk:

$$E_{\theta_1} S_\tau^r = E_{\theta_0}[S_\tau^r \exp(\xi_1 S_\tau); \tau < \infty]$$

$$(5.14) \quad = E[Y^r \exp(\xi_1 Y); T < \infty].$$

Here  $T$  is the epoch of first exceedance of, and  $Y = S_T - \epsilon T$  is the overshoot of, the linear boundary with slope  $\epsilon$  passing through the origin. The limiting relation

$$Y^r \exp(\xi_1 Y) I_{\{T < \infty\}} \rightarrow S_\tau^r$$

clearly holds for (almost) every sample point, so it suffices to show that the function

$$E[Y^{2r} \exp(2\xi_1 Y); T < \infty] = E_{\theta_0}[S_\tau^{2r} \exp(2\xi_1 S_\tau); \tau < \infty]$$

of  $\epsilon$  is bounded for small  $\epsilon$ .

For this we use Lemma 2 below. From assumption 0 it follows that  $EX^- > 0$ . Choose  $c \in (0, \infty)$  large enough that  $E(X^- \wedge c) > 0$ . Apply Lemma 2 with  $f(x) = x^{2r} \exp(2\xi_1 x)$  to conclude

$$(5.15) \quad 0 \leq E_{\theta_0}[S_\tau^{2r} \exp(2\xi_1 S_\tau); \tau < \infty] \leq E_{\theta_0}[(X^+)^{2r} \exp(2\xi_1 X^+) (X^+ + c)] / E_{\theta_0}(X^- \wedge c).$$

As  $\epsilon \downarrow 0$ , the last expression in (5.15) will by the dominated convergence theorem approach

$$E[(X^+)^{2r} (X^+ + c)] / E(X^- \wedge c) < \infty. \quad \blacksquare$$

Let  $(S_n)_{n \geq 0}$  be any random walk with  $\liminf S_n < +\infty$ . (When  $EX \in [-\infty, \infty]$  exists this is equivalent to  $EX \leq 0$ , so our given  $(S_n)$  is such a walk under  $P_{\theta_0}$ .) Lemma 2 provides simple upper bounds on certain functionals of the first ladder height  $S_\tau$ . In the course of proving Lemma 2 we shall make use of the observation that if  $(U(t)) = (EN(t))$  is the renewal function for a persistent renewal process  $N$  (convention:  $N(0) = 1$ ) with interevent distribution  $\mathcal{L}(\xi)$ , then for any  $t \geq 0$  and  $c > 0$

$$(5.16) \quad U(t) \leq (t + c) / E(\xi \wedge c).$$

Indeed, if we suppose at first that  $\xi \leq c$ , then clearly  $S_{N(t)} \leq t + c$ , and (5.16) follows from Wald's identity  $ES_{N(t)} = (E\xi) \times U(t)$ . The boundedness condition on  $\xi$  can then be removed by a simple truncation argument. When  $N$  is the renewal process corresponding to (the magnitudes

of) the (weak) descending ladder heights, (5.16) yields

$$(5.17) \quad U(t) \leq (t + c)/E(X^- \wedge c)$$

since

$$E(|S_{\tau_-}| \wedge c) \geq E(|S_{\tau_-}| \wedge c; \tau_- = 1) = E(X^- \wedge c);$$

here  $\tau_- = \inf\{n \geq 1 : S_n \leq 0\}$  is the (almost surely finite) first descending ladder epoch.

**Lemma 2.** *Let  $(S_n)_{n \geq 0}$  be any random walk with  $\liminf S_n < +\infty$ . If  $f : [0, \infty) \rightarrow [0, \infty)$  is measurable and nondecreasing, then for any  $c > 0$*

$$(5.18) \quad E[f(S_\tau); \tau < \infty] \leq E[f(X)(X + c); X > 0]/E(X^- \wedge c).$$

*Proof.* Since  $S_\tau > 0$  over  $\{\tau < \infty\}$ , we may suppose  $f(0) = 0$  and establish (5.18) in the form

$$(5.19) \quad E[f(S_\tau); \tau < \infty] \leq E[f(X^+)(X^+ + c)]/E(X^- \wedge c).$$

In light of (5.17) it is enough to show

$$(5.20) \quad E[f(S_\tau); \tau < \infty] \leq E[f(X^+)U(X^+)];$$

here  $U$  is the distribution function for the renewal measure

$$(5.21) \quad U(dy) = \sum_{k=0}^{\infty} P\{|S_{\tau^{(k)}}| \in dy\}, \quad y \in [0, \infty),$$

$\tau^{(k)}$  denoting the  $k$ th descending ladder epoch. The duality result

$$(5.22) \quad P\{S_\tau \in B; \tau < \infty\} = \int_{[0, \infty)} U(dy)P\{X \in B + y\}$$

is well-known; it can be proved by employing proposition 8.38 and the version – valid for any random walk –

$$(5.23) \quad P\{S_{\tau_+} > x; \tau_+ < \infty\} = \sum_{n=0}^{\infty} \int_{(-\infty, 0]} P\{\tau_+ > n, S_n \in dy\}P\{X > x - y\}$$

of (8.85) in Siegmund (1985). From (5.22) follows

$$\begin{aligned}
 E[f(S_\tau); \tau < \infty] &= \int_{[0, \infty)} \int_{(0, \infty)} f(s) P\{X - y \in ds\} U(dy) \\
 &= \int_{[0, \infty)} E f((X - y)^+) U(dy) \quad (\text{recall } f(0) = 0) \\
 &= E \int_{[0, \infty)} f((X^+ - y)^+) U(dy) \\
 &= E \int_{[0, X^+]} f(X^+ - y) U(dy) \\
 &\leq E[f(X^+) U(X^+)],
 \end{aligned}$$

the inequality holding because  $f$  is nondecreasing. ■

**5.2. Small- $\epsilon$  analysis of  $(\epsilon_m)$  and  $(\lambda_m)$ .** Throughout Section 5.2 we suppose that assumptions 0 and 2b are met and that  $0 \in I^0$ . We shall spend little effort examining the sequence  $(\epsilon_m)$  since its computation via (1.14)–(1.15) is quite simple. The following two remarks shall suffice. (i) For fixed  $m$ , as  $\epsilon \downarrow 0$  the quantity  $h^{-1}(g(m) - bm)$  decreases to  $h^{-1}(a - bm)$ ; it follows that  $\delta_m = \delta_m(\epsilon)$ , and hence also the ratio of  $\epsilon_m = \epsilon_m(\epsilon)$  to  $\epsilon h$  converges to  $\delta_m(0) :=$  fractional part of  $h^{-1}(a - bm)$ . Note, however, that the convergence of  $\epsilon_m/(\epsilon h)$  to  $\delta_m(0)$  is not uniform in  $m$ . (ii) If  $V = V(\epsilon)$  is a random variable having the same distribution as the sequence  $(\epsilon_m)$ , then as  $\epsilon \downarrow 0$  the random variables  $(\epsilon h)^{-1}V(\epsilon)$  converge in distribution to the uniform distribution on  $[0, 1]$ . In this sense when  $\epsilon$  is small the values  $\epsilon_m$  are approximately equidistributed over  $[0, \epsilon h]$ . Correspondingly, the expression  $(|\theta_0|C)^{-1} - 1$  in (3.2) equals  $(1 + o(1))\epsilon h/2$  as  $\epsilon \downarrow 0$ .

The remainder of this section is devoted to a proof of the limit result (1.26) for  $(\lambda_m)$ , which implies that if  $V^* = V^*(\epsilon)$  is a random variable having the same distribution as the sequence  $(\lambda_m)$ , then

$$(5.24) \quad (\epsilon h)^{-1}V^*(\epsilon) \rightarrow 1/2 \quad \text{in probability as } \epsilon \downarrow 0.$$

To prove (1.26) it clearly suffices by (3.3) to establish

$$(5.25) \quad E(\delta_m + rW) \rightarrow 1/2 \quad \text{uniformly in } m \text{ as } \epsilon \downarrow 0,$$

with  $r = r(\epsilon) = h^{-1}(\epsilon - b)$ , and  $W$  and  $\langle x \rangle$  as defined in Section 3 above. Because  $\langle \delta_m + rW \rangle =$

$\langle \delta_m + \langle rW \rangle \rangle$  the proof of (5.25) is accomplished by means of the following two lemmas, in which  $U$  denotes a random variable uniformly distributed on  $[0, 1]$ .

**Lemma 3.** As  $\epsilon \downarrow 0$ ,

$$(5.26) \quad \langle rW \rangle \rightarrow U \text{ in distribution.}$$

**Lemma 4.** If  $U_n \in [0, 1]$  are random variables converging in distribution to  $U$  as  $n \rightarrow \infty$ , then as  $n \rightarrow \infty$

$$(5.27) \quad E\langle x + U_n \rangle \rightarrow E\langle x + U \rangle = 1/2 \text{ uniformly in } x \in [0, 1].$$

*Proof of Lemma 3.* The assertion (5.26) is equivalent to the assertion that for  $j \neq 0$  the Fourier coefficient

$$(5.28) \quad E \exp(2\pi i j r W) = \sum_{n=0}^{\infty} \epsilon^{2\pi i j r n} u_n / \tilde{u}(1) = \tilde{u}(\epsilon^{2\pi i j r}) / \tilde{u}(1)$$

vanishes in the limit as  $\epsilon \downarrow 0$ . According to (1.17),

$$(5.29) \quad \begin{aligned} \tilde{u}(\epsilon^{2\pi i j r}) / \tilde{u}(1) &= \exp \left[ -\frac{1}{2} \sum_{k=1}^{\infty} k^{-1} \exp(-k\phi(\theta_0)) \right] \\ &\times \exp \left[ \frac{1}{2} \sum_{k=1}^{\infty} \epsilon^{2\pi i j k r} k^{-1} \exp(-k\phi(\theta_0)) \right] \\ &\times \exp \left[ \sum_{k=1}^{\infty} (\epsilon^{2\pi i j k r} - 1) k^{-3/2} \exp(-k\phi(\theta_0)) k^{1/2} (P_0\{S_k > 0\} - 1/2) \right]. \end{aligned}$$

In examining the first two factors on the right in (5.29) we shall make use of the well-known fact that, for  $|z| < 1$ ,  $\sum_{k=1}^{\infty} k^{-1} z^k = -L(1-z)$ , where  $L$  denotes the principal branch of the logarithm function. Thus

$$(5.30) \quad \exp \left[ -\frac{1}{2} \sum_{k=1}^{\infty} k^{-1} \exp(-k\phi(\theta_0)) \right] = \exp \left[ \frac{1}{2} L(1 - e^{-\phi(\theta_0)}) \right] \sim 2^{-1/2} \epsilon$$

as  $\epsilon \downarrow 0$  and

$$(5.31) \quad \exp \left[ \frac{1}{2} \sum_{k=1}^{\infty} \epsilon^{2\pi i j k r} k^{-1} \exp(-k\phi(\theta_0)) \right] = \exp \left[ -\frac{1}{2} L(1 - e^{2\pi i j r} e^{-\phi(\theta_0)}) \right].$$

For (5.31) we consider two possibilities according to the value of  $jr(0)$ , where

$$(5.32) \quad r(0) := \lim_{\epsilon \downarrow 0} r(\epsilon) = -h^{-1}b \in (-1, 0].$$

- (a) If  $jr(0) \notin \mathbb{Z}$ , then (5.31) tends to  $\exp[-\frac{1}{2}L(1 - \exp(2\pi i jr(0)))] \in \mathbb{C}$  as  $\epsilon \downarrow 0$ .
- (b) If  $jr(0) \in \mathbb{Z}$ , then  $jr$  and  $j\epsilon h^{-1}$  differ by an integer and  $\exp[-\frac{1}{2}L(1 - e^{2\pi i jr} e^{-\phi(\theta_0)})] \sim h^{1/2}(2\pi j)^{-1/2} \exp(\pi i/4)\epsilon^{-1/2}$  as  $\epsilon \downarrow 0$ . In either case the product of (5.30) and (5.31) is  $O(\epsilon^{1/2}) = o(1)$ .

Regarding the third factor on the right in (5.29), by the Berry-Esséen theorem (e.g., Feller, 1971, theorem XVI.5.1)

$$(5.33) \quad |(\epsilon^{2\pi i j k r} - 1)k^{-3/2} \exp(-k\phi(\theta_0))k^{1/2}(P_0\{S_k > 0\} - 1/2)| \leq 2Ck^{-3/2}\sigma^{-3}E_0|X|^3$$

where  $C$  is a universal constant. In light of the facts  $\sigma^2 \rightarrow 1$  and  $E_0|X|^3 \rightarrow E|X|^3$ , it is apparent that the third factor on the right in (5.29) remains bounded for all small  $\epsilon$ . We conclude

$$(5.34) \quad \tilde{u}(\epsilon^{2\pi i j r})/\tilde{u}(1) = O(\epsilon^{1/2}) = o(1) \quad \text{as } \epsilon \downarrow 0.$$

thereby completing the proof of Lemma 3. ■

*Proof of Lemma 4.* The difference RHS - LHS in (5.27) equals

$$(5.35) \quad (EU_n - EU) + (E[x + U] - E[x + U_n]).$$

where  $[x]$  denotes the integer part of  $x$ , i.e., the greatest integer no larger than  $x$ . The first of the two differences in (5.35) is independent of  $x$  and tends to zero. The second equals  $P\{U \geq 1 - x\} - P\{U_n \geq 1 - x\}$  and tends to zero uniformly in  $x \in [0, 1)$  by a theorem of Pólya (e.g., Chung, 1974, exercise 4.3.4). ■

**Acknowledgments.** The authors thank Professor David Siegmund for several useful comments.

### **References**

1. Bahadur, R. R. and Rao, R. Ranga: On deviations of the sample mean. *Ann. Math. Statist.* **31**, 1015-1027 (1960)
2. Blackwell, D. and Hodges, J. L., Jr.: The probability in the extreme tail of a convolution. *Ann. Math. Statist.* **30**, 1113-1120 (1959)
3. Chung, K. L.: *A Course in Probability Theory*, 2nd ed. New York: Academic Press 1974
4. Feller, W.: *An Introduction to Probability Theory and Its Applications*, Vol. 2, 2nd ed. New York: Wiley 1971
5. Fill, J. A.: Convergence rates related to the strong law of large numbers. Stanford University Technical Report No. 14 (1980)
6. Fill, J. A.: Convergence rates related to the strong law of large numbers. *Ann. Probability* **11**, 123-142 (1983)
7. Fill, J.A.: Asymptotic expansions for large deviation probabilities in the strong law of large numbers. University of Chicago Technical Report No. 215 (1987)
8. Lai, T.Z.: Asymptotic moments of random walks with applications to ladder variables and renewal theory. *Ann. Probability* **4**, 51-66 (1976)
9. Siegmund, D.: Large deviation probabilities in the strong law of large numbers. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **31**, 107-113 (1975)
10. Siegmund, D: Sequential Analysis. New York: Springer-Verlag 1985
11. Spitzer, F.: A Tauberian theorem and its probability interpretation. *Trans. Amer. Math. Soc.* **94**, 150-169 (1960)

~~UNCLASSIFIED~~  
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

A187 568

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 3	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle)  THE CONVERGENCE RATE FOR THE STRONG LAW OF LARGE NUMBERS: GENERAL LATTICE DISTRIBUTIONS		5. TYPE OF REPORT & PERIOD COVERED  Technical Report
7. AUTHOR(s)  James Allen Fill and Michael J. Wichura		6. PERFORMING ORG. REPORT NUMBER  N00014-87-K-0078
9. PERFORMING ORGANIZATION NAME AND ADDRESS  Department of Statistics - Sequoia Hall Stanford University Stanford, California 94305-4065		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS  HR-042-373
11. CONTROLLING OFFICE NAME AND ADDRESS  Statistics & Probability Program Code (411 (SP)) Office of Naval Research Arlington, Virginia 22217		12. REPORT DATE  October 1987
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		13. NUMBER OF PAGES  30pp
		15. SECURITY CLASS. (of this report)  Unclassified
		16a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)  Random walk; laws of large numbers; convergence rates; boundary crossing probabilities; large deviations; lattice distribution; associated distributions; renewal theory.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  See abstract on back.		

20. ABSTRACT.

Let  $X_1, X_2, \dots$  be a sequence of independent random variables with common lattice distribution function  $F$  having zero mean, and let  $(S_n)$  be the random walk of partial sums. The strong law of large numbers (SLLN) implies that for any  $\alpha \in \mathbb{R}$  and  $\epsilon > 0$

$$p_m := P\{S_n > \alpha + \epsilon n \text{ for some } n \geq m\}$$

decreases to 0 as  $m$  increases to  $\infty$ . Under conditions on the moment generating function of  $F$ , we obtain the convergence rate by determining  $p_m$  up to asymptotic equivalence. When  $\alpha = 0$  and  $\epsilon$  is a point in the lattice for  $F$ , the result is due to Siegmund [Z. Wahrscheinlichkeitstheorie verw. Gebiete 31 (1975):107–113]; but this restriction on  $\epsilon$  precludes all small values of  $\epsilon$ , and these values are the most interesting vis-à-vis the SLLN. Even when  $\alpha = 0$  our result handles important distributions  $F$  for which Siegmund's result is vacuous, for example, the two-point distribution  $F$  giving rise to simple symmetric random walk on the integers.

END

FEB.

1988

DTIC